U-Net, U-Net++ and PSPNet

Cell image segmentation using deep learning

Mateusz Chojnacki, Ignacy Makowski, Younginn Park, Lidia Stadnik

University of Warsaw

Aim of the project

3 pretrained models from Segmentation Models in Pytorch library were selected (U-Net, U-Net++ and PSPNet, with pretrained weights on ImageNet with ResNet34 as a backbone architecture). In addition, a Swin-Unet model was included, trained from scratch without pretrained weights.

Perform fine-tuning on our cell microscopy training set using these 4 models.

Evaluate the performance of these models on the task of semantic segmentation on the Cell Segmentation dataset from NeurIPS 2022 challenge.

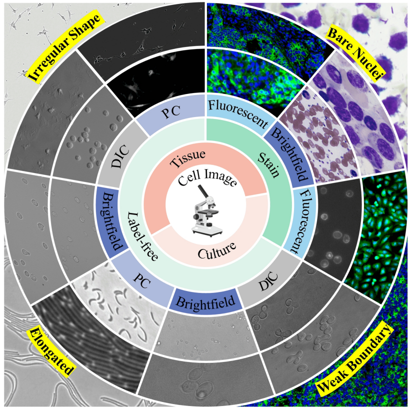
NeurIPS 2022 dataset

From 2022 cell segmentation challenge from NeurIPS

Consist of three part, containing different cell images in: tiff, tif, png, jpg and bmp formats.

Due to task description, we selected only grayscale .tiff and .tif images

1. Training - we limit ourselves to the labeled subset
    - original size: 1000,
    - after preprocessing: **491,** from which **20** were used as a validation dataset
2. Tuning (used as a test dataset)
    - original size: 100,
    - after preprocessing: **58**
3. Testing (skipped)
    - huge and impractical to use



https://neurips22-cellseg.grand-challenge.org/dataset/

U-Net

Type of convolutional neural network architecture containing two major parts:

- contracting path - decreasing size of data matrix due to usage of pooling layers and convolutional layers
- expansive path - increasing size of data matrix using up-convolution (type of interpolation)

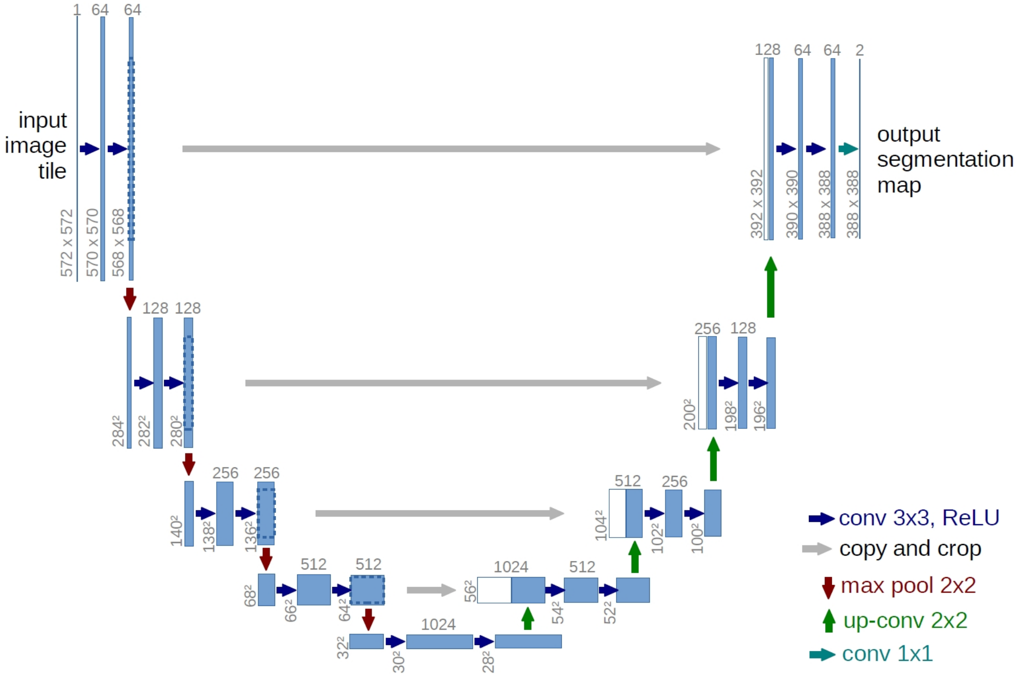
main idea - skip-connection enabling flow of data from contracting path to expansive path

Home > Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 > Conference paper

U-Net: Convolutional Networks for Biomedical Image Segmentation

Conference paper | First Online: 18 November 2015

pp 234–241 | Cite this conference paper



U-Net++

Electrical Engineering and Systems Science > Image and Video Processing

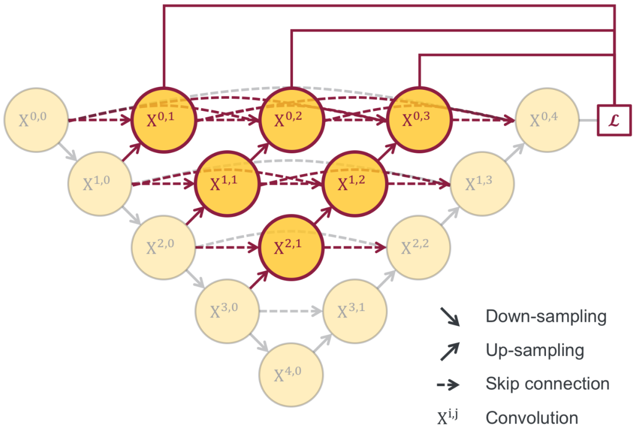
[Submitted on 11 Dec 2019 (v1), last revised 28 Jan 2020 (this version, v2)]

UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation

Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, Jianming Liang

This type of convolutional network model architecture built on idea used in U-net, but with three differences:

- using several U-nets of various depth, which share encoder and are trained simultaneously to alleviate problem of unknown 'true' depth of ideal model.
- improving skip-connections to include multiple sub-U-nets data (fully connected in each level).
- algorithm for pruning much more complicated model, to fasten inference



https://github.com/MrGiovanni/UNetPlusPlus/blob/master/Figures/fig_unet++.png

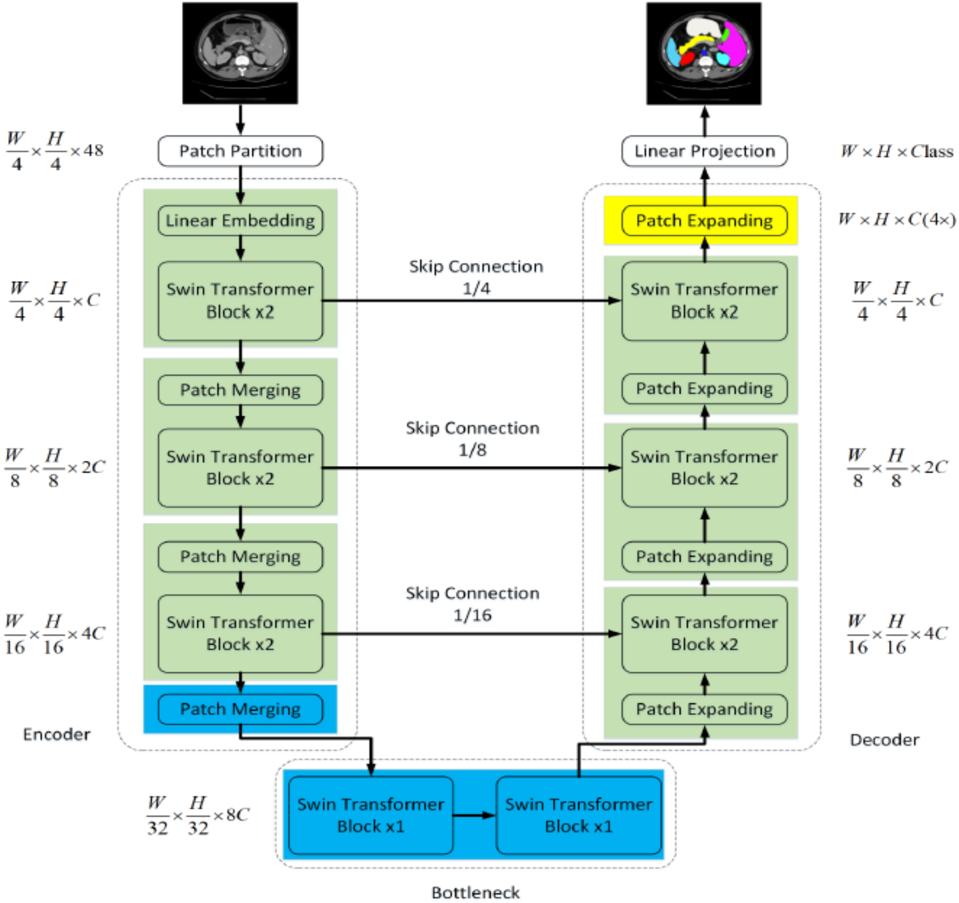
Swin U-Net

Electrical Engineering and Systems Science > Image and Video Processing

[Submitted on 12 May 2021]

Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation

Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian, Manning Wang



Swin-Unet integrates Swin Transformers into a U-Net-like encoder-decoder structure for medical image segmentation:

Encoder:

- Divides input into non-overlapping patches.
- Processes through hierarchical Swin Transformer blocks.
- Performs patch merging for downsampling.

Bottleneck:

- Deepest layer, with powerful Swin blocks for contextual encoding.

Decoder:

- Patch expanding gradually upsamples features.
- Combines with encoder outputs via skip connections to refine predictions.

Output:

- Final linear projection creates pixel-wise segmentation map.

PSPNet

Computer Science > Computer Vision and Pattern Recognition

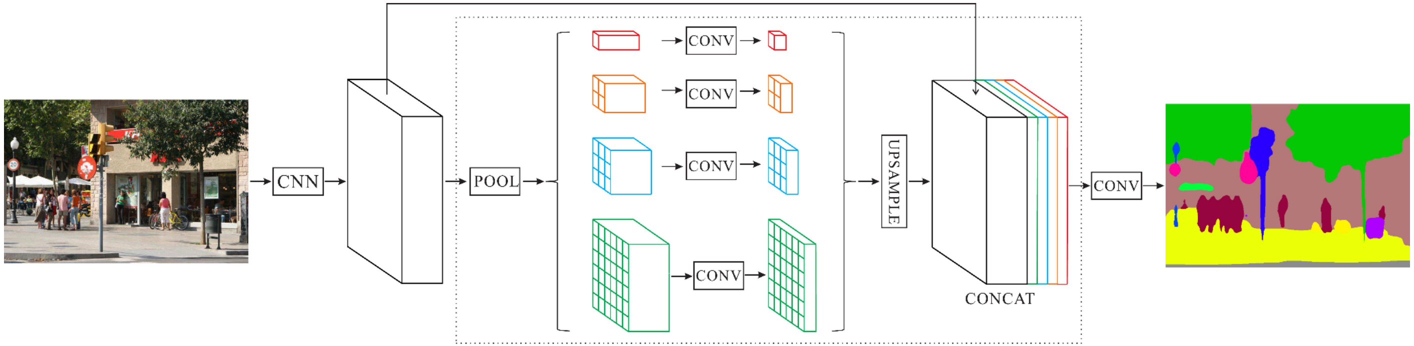
[Submitted on 4 Dec 2016 (v1), last revised 27 Apr 2017 (this version, v2)]

Pyramid Scene Parsing Network

Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, Jiaya Jia

This type of convolutional network model architecture uses both skip-connection as well as pyramid pooling module, which acts as a global prior. PPM works to counteract loss of information from receptive field in majority cases during convolution, when picture is much smaller than input receptive field.

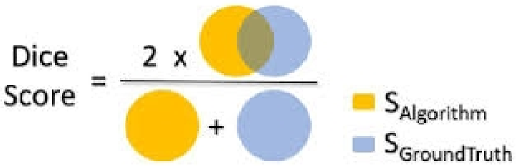
Architecture achieves this, by performing convolution on 4 different scales of whole image, ½ and small parts simultaneously.

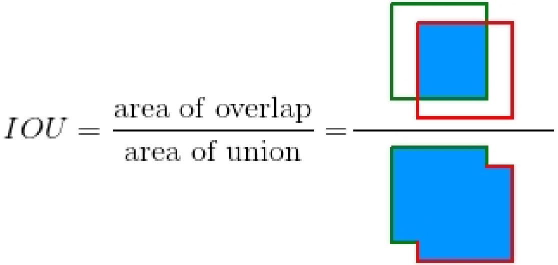


(a) Input Image (b) Feature Map (c) Pyramid Pooling Module (d) Final Prediction

https://paperswithcode.com/method/pspnet

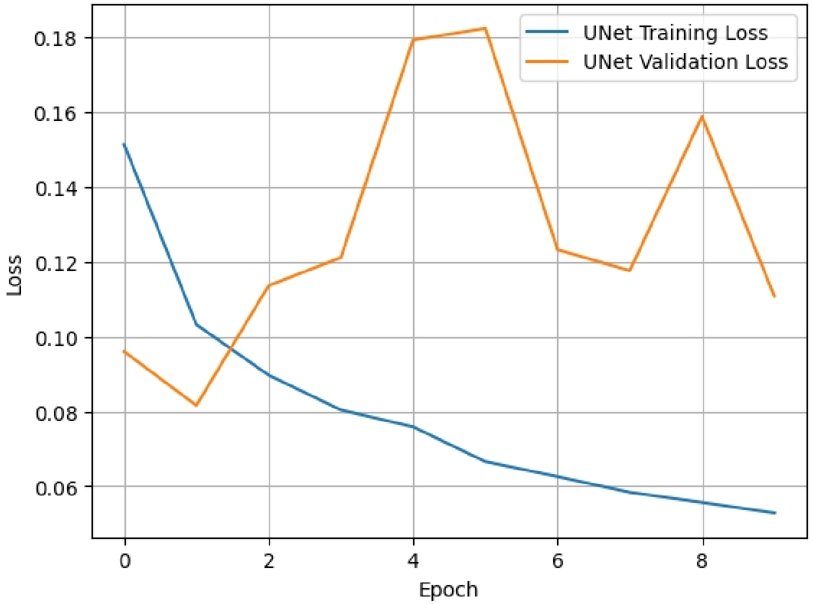
Model performance





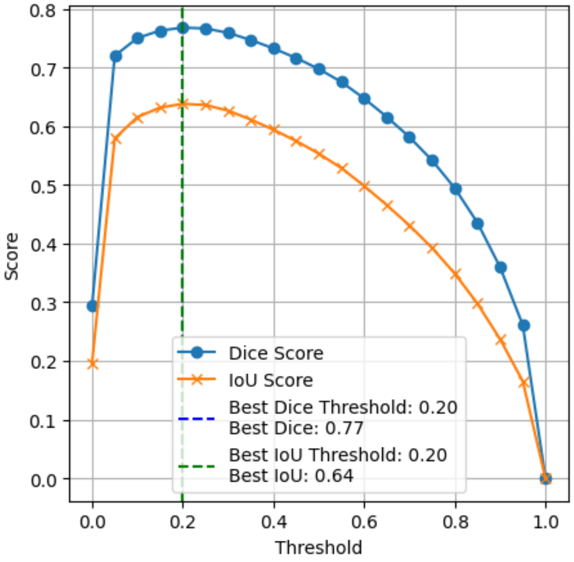
Model	Best Threshold	Dice	IoU
base U-Net1 (de novo, basic architecture)	-	0.34	0.22
base U-Net2 (de novo, with BatchNorm)	-	0.41	0.30
U-Net (fine-tuned, ImageNet pretrained)	0.20	0.77	0.64
U-Net (fine-tuned, ImageNet pretrained + augmentation 1 & 2)	0.5 (two classes)	**0.93** **0.89**	**0.86** **0.80**
U-Net++ (fine-tuned, ImageNet pretrained)	0.05	0.38	0.26
PSPNet (fine-tuned, ImageNet pretrained)	0.05	0.61	0.42
PSPNet(fine-tuned, ImagNet pretrained + augmentation)	(two class)	**0.86**	**0.75**
Swin U-Net (fine-tuned + augmentation)	0.5	**0.81**	**0.71**

Training U-net (without augmentation)

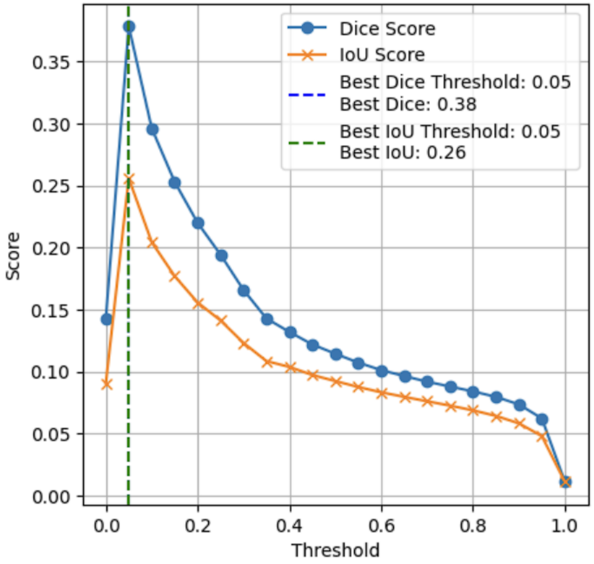


Model training and validation losses

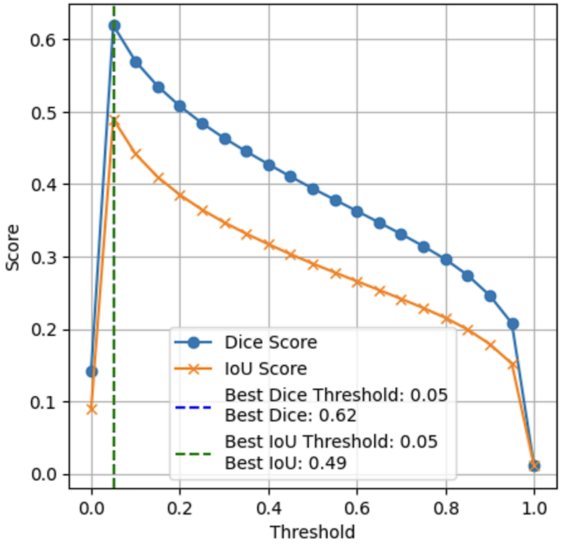
Threshold optimization



UNet

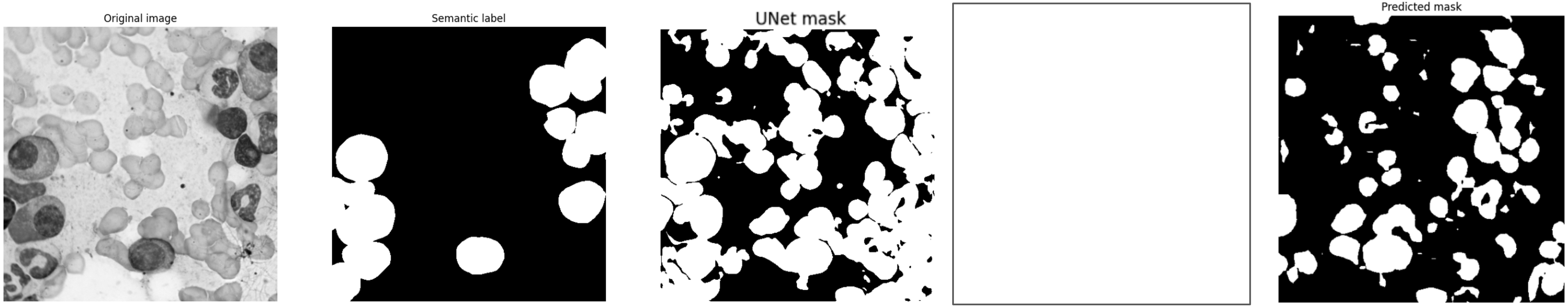


UNet++



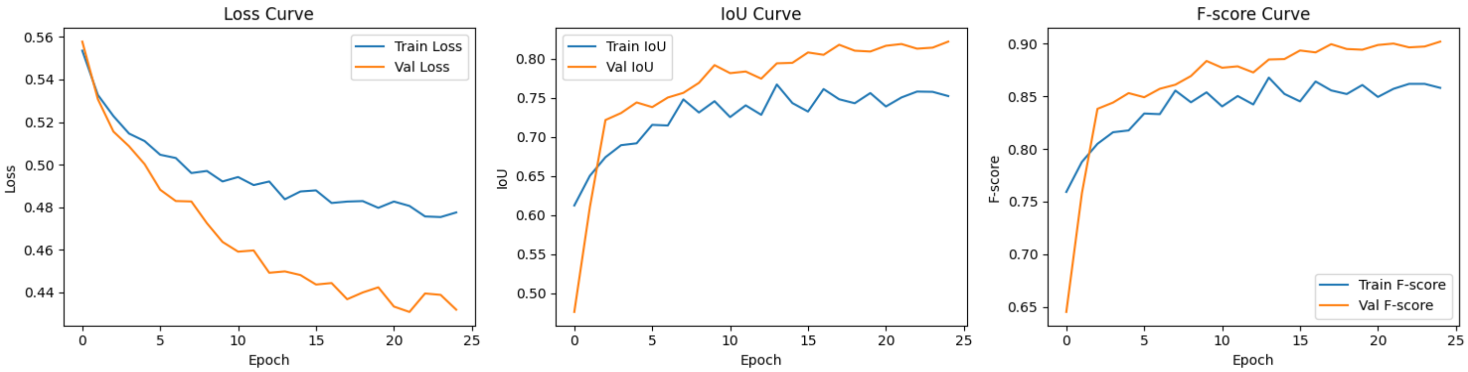
PSPNet

Segmentation mask comparison

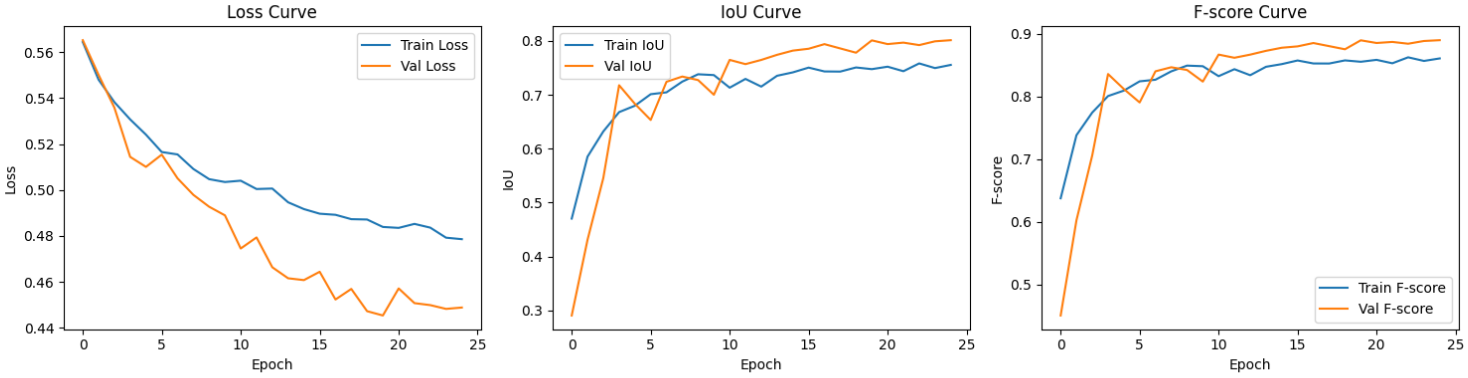


Examples of segmentation masks predicted by selected models

Training Unet with augmented data

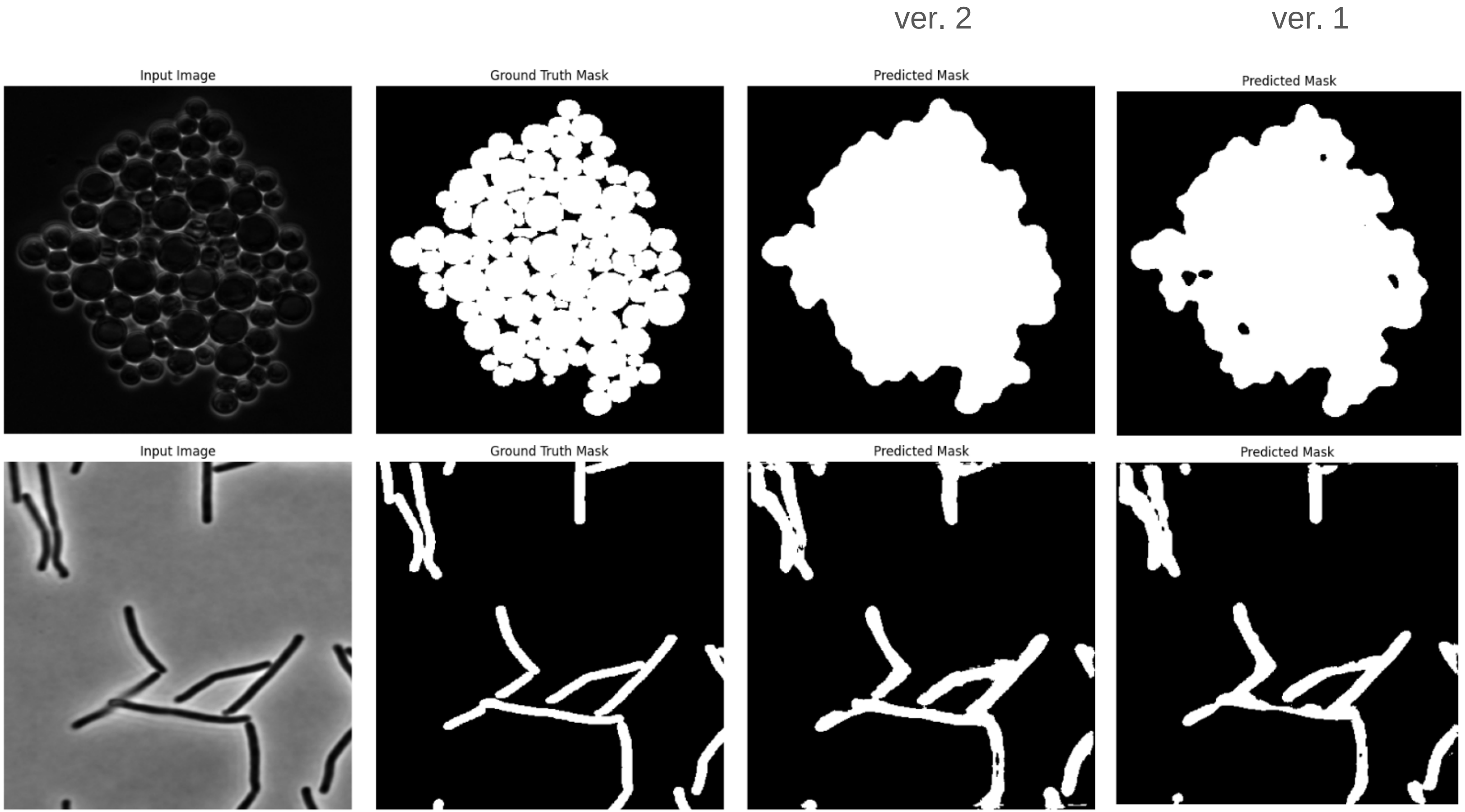


version 1

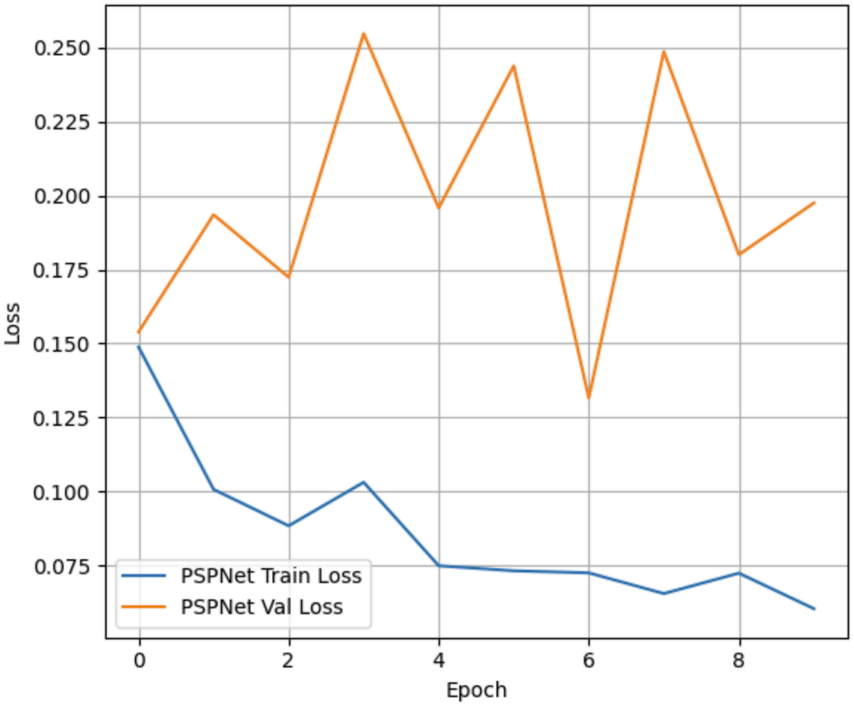


version 2

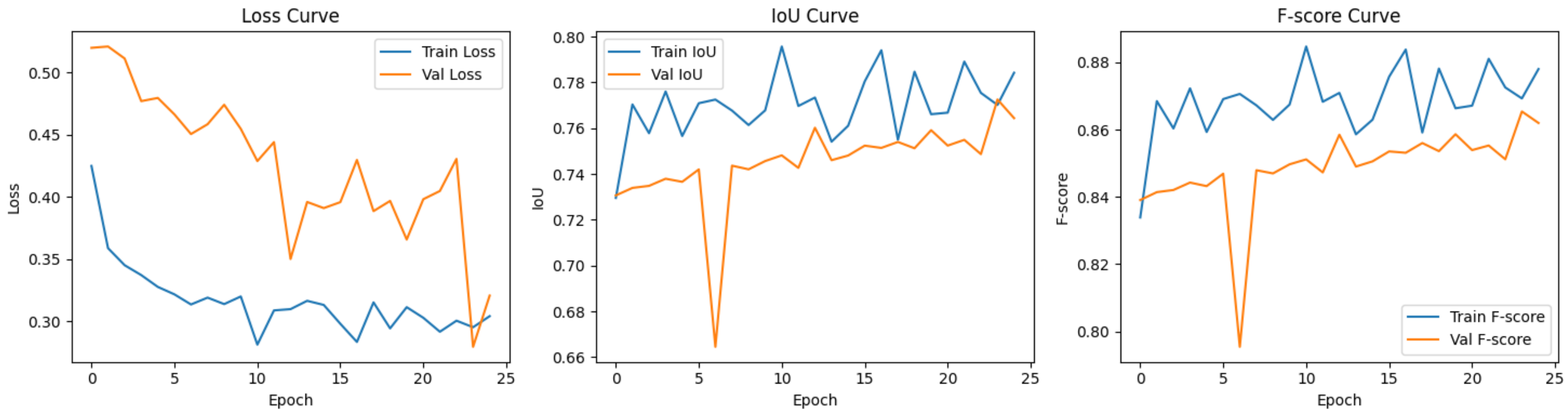
Unet with augmented data



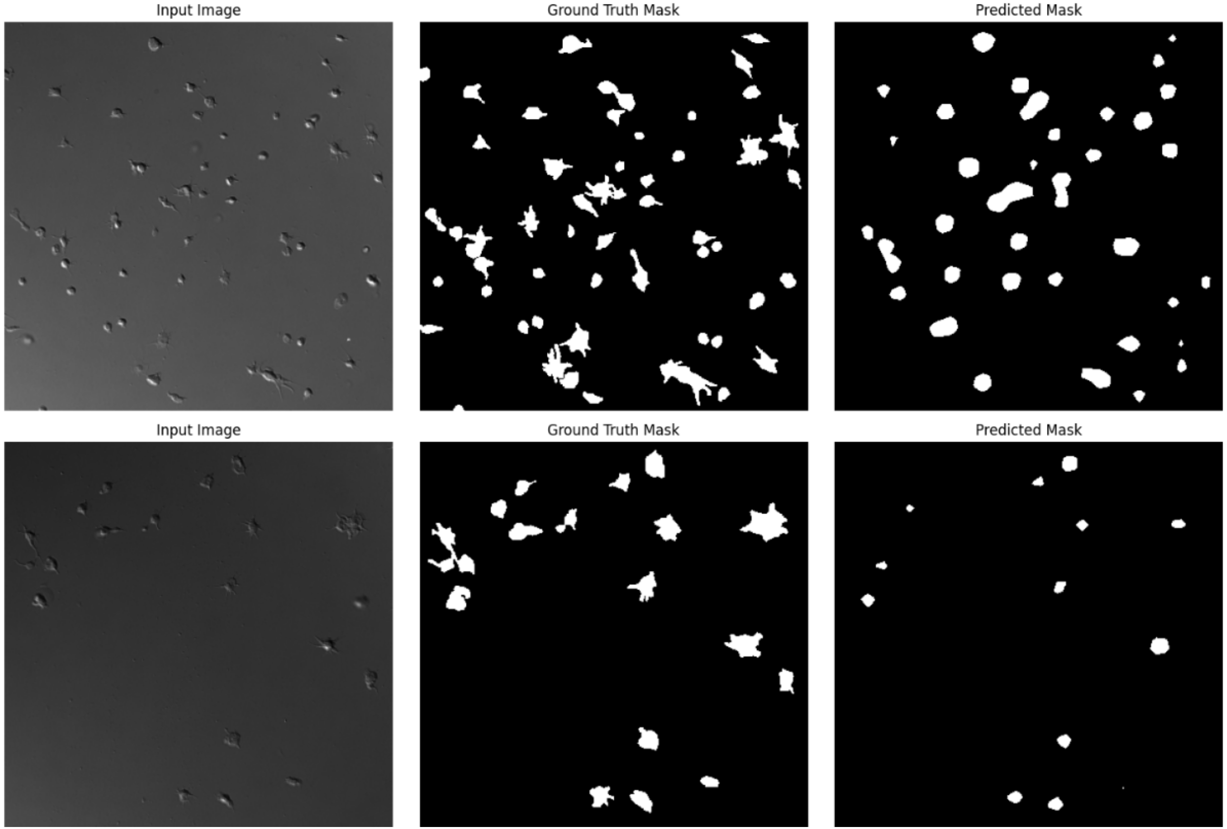
Training PSPNet without augmentation



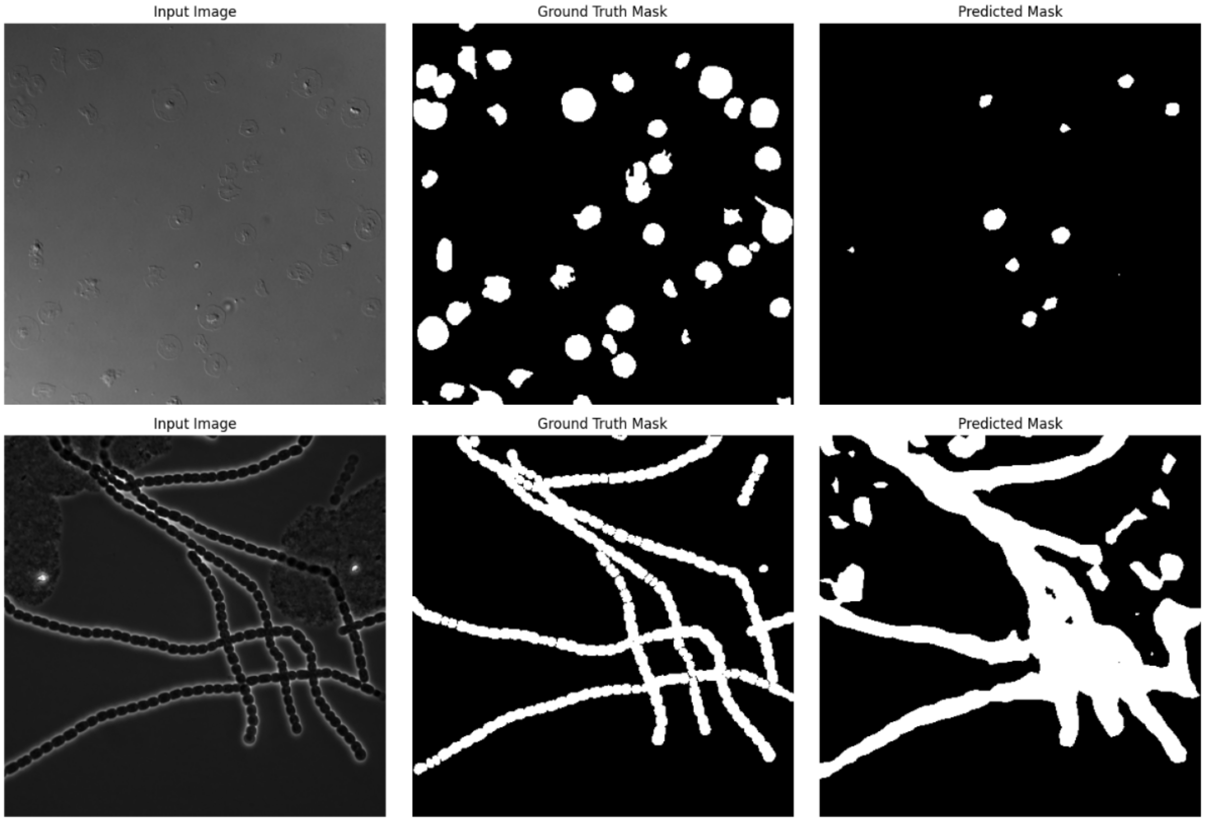
Training PSPNet with augmented data



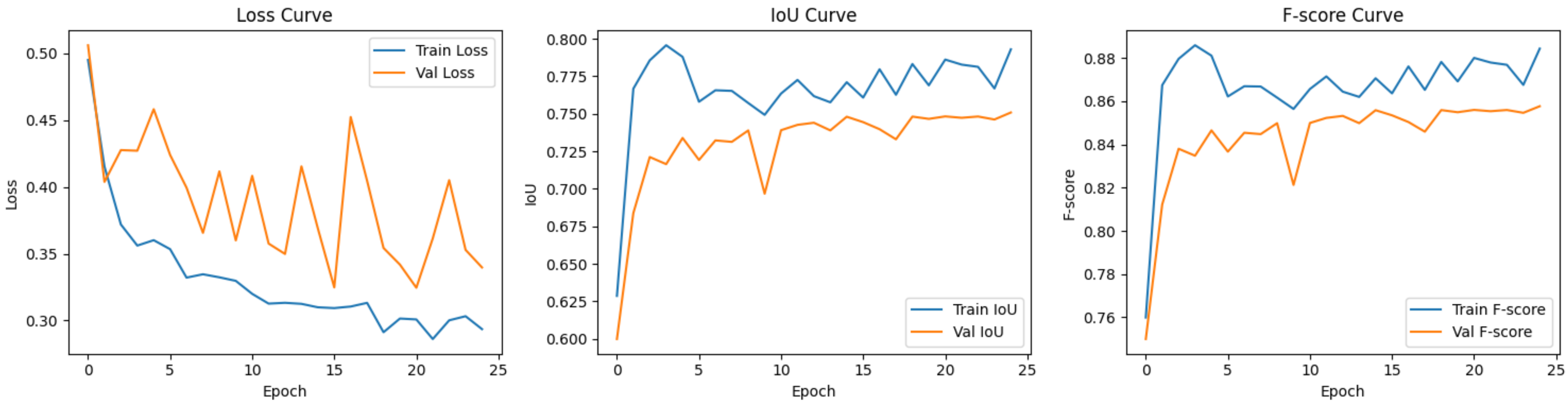
Result: PSPNet with augmented data



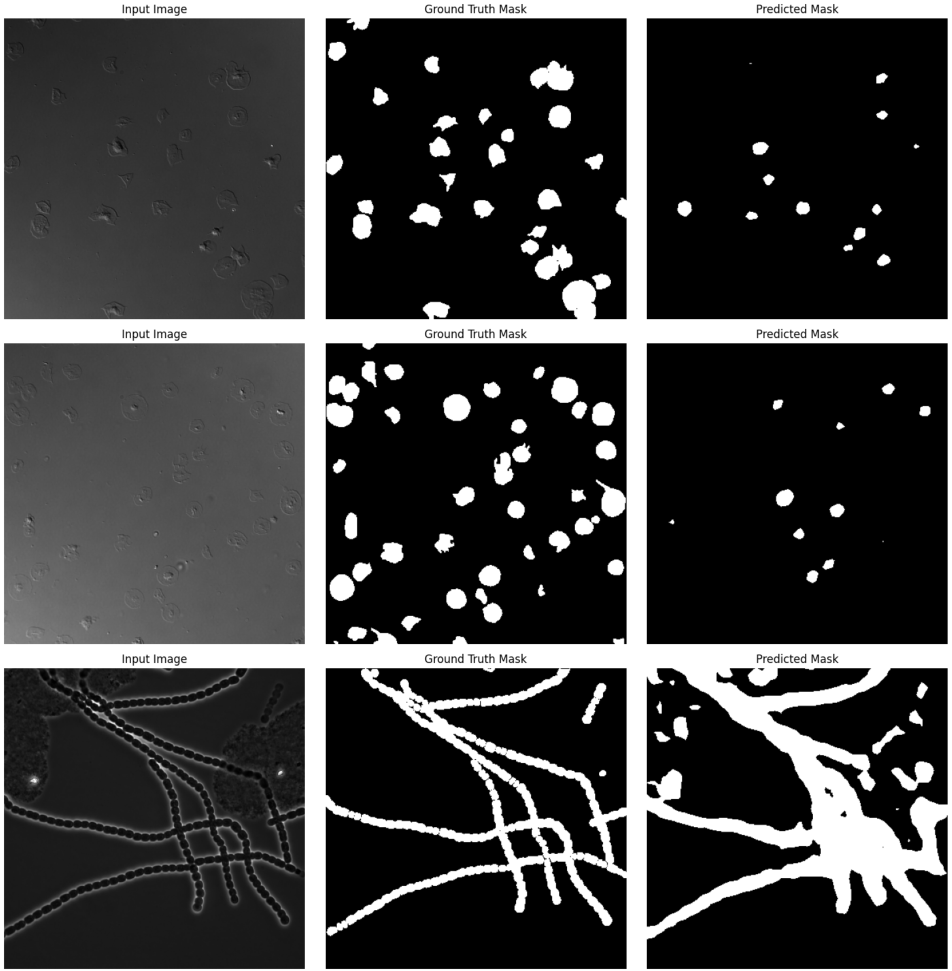
Result: PSPNet with augmented data

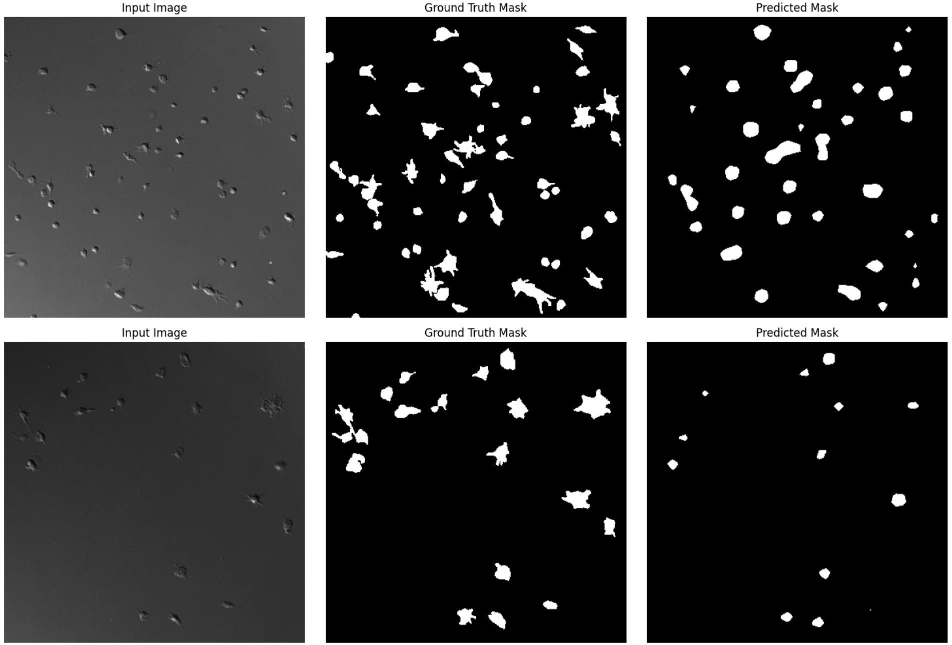


Training Swin U-Net with augmentation



Result: Swin U-net with augmented data





References

- **Dataset** - Ma, J., Xie, R., Ayyadhury, S., Ge, C., Gupta, A., Gupta, R., Gu, S., Zhang, Y., Lee, G., Kim, J., Lou, W., Li, H., Upschulte, E., Dickscheid, T., de Almeida, J. G., Wang, Y., Han, L., Yang, X., Labagnara, M., … Wang, B. (2024). NeurIPS 2022 Cell Segmentation Competition Dataset [Data set]. Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS), New Orleans. Zenodo. https://doi.org/10.5281/zenodo.10719375
- **U-Net** - O. Ronneberger and P.Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", Medical Image Computing and Computer-Assisted Intervention (MICCAI), LNCS, 9351, 234--241, 2015, Springer, http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a
- **U-Net++** - Zongwei Zhou and Md Mahfuzur Rahman Siddiquee and Nima Tajbakhsh and Jianming Liang, UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation, 2020, 1912.05074, hhttps://arxiv.org/abs/1912.05074
- **PSPNet** - Zhao, Hengshuang and Shi, Jianping and Qi, Xiaojuan and Wang, Xiaogang and Jia, Jiaya, Pyramid Scene Parsing Network, CVPR, 2017
- **Swin U-Net** - Hatamizadeh et al., 2022. Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation
- **IoU** - Padilla, Rafael & Netto, Sergio & da Silva, Eduardo. (2020). A Survey on Performance Metrics for Object-Detection Algorithms. 10.1109/IWSSIP48289.2020.
- **Dice** - Rao, Divya & K, Prakashini & Singh, Rohit & J, Vijayananda. (2022). Automated segmentation of the larynx on computed tomography images: a review. Biomedical Engineering Letters. 12. 1-9. 10.1007/s13534-022-00221-3.
- **Segmentation Models in PyTorch** - Iakubovskii, Pavel. (2019). Segmentation Models Pytorch. GitHub repository. GitHub. https://github.com/qubvel/segmentation_models.pytorch